JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2019, 15, 3694−3709

pubs.acs.org/JCTC

Massively Parallel Implementation of Steered Molecular Dynamics in Tinker-HP: Comparisons of Polarizable and Non-Polarizable Simulations of Realistic Systems

Frédéric Célerse,[†,‡] Louis Lagardère,[¶,§,∥] Etienne Derat,[‡] and Jean-Philip Piquemal*[,†,⊥,#]

[†]Laboratoire de Chimie Théorique, UMR 7616 CNRS, Sorbonne Université, 75005 Paris, France

[‡]Institut Parisien de Chimie Moléculaire, UMR 8232 CNRS, Sorbonne Université, 75005 Paris, France

[¶]Institut des Sciences du Calcul et des Données, Sorbonne Université, 75005 Paris, France

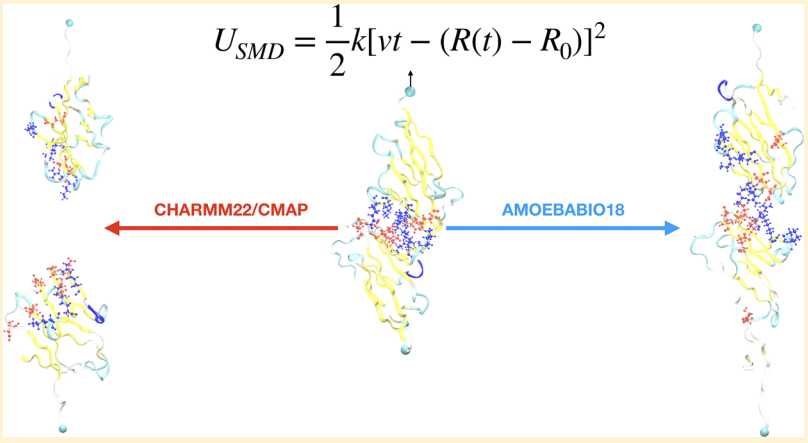
[§]Institut Parisien de Chimie Physique et Théorique, FR 2622 CNRS, Sorbonne Université, 75005 Paris, France

[∥]Laboratoire de Chimie théorique, UMR 7616 CNRS, Sorbonne Université, 75005 Paris, France

[⊥]Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas 78712, United States

[#]Institut Universitaire de France, 75005 Paris, France

Supporting Information



ABSTRACT: Steered molecular dynamic (SMD) is a powerful technique able to accelerate rare event sampling in Molecular Dynamics (MD) simulations by applying an external force to a set of chosen atoms. Despite generating nonequilibrium simulations, SMD remains capable of reconstructing equilibrium properties such as the Potential of Mean Force (PMF). Of course, one would like to use all types of force fields (FF) ranging from classical ones to more advanced polarizable models using point induced dipoles and distributed multipoles such as AMOEBA. To enable such studies, the SMD methodology has been implemented in the framework of the massively parallel Tinker-HP software allowing for both long polarizable and non-polarizable MD simulations of large proteins. To validate this new implementation, we first compared the Tinker-HP SMD results to the literature. Tests have been performed on three different benchmark systems: the M−A deca-alanine (112 atoms), the ubiquitin (9737 atoms), and the CD2CD58 complex (97594 atoms). Non-polarizable (AMBER99, AMBER99SB, CHARMM22/CMAP, and OPLS−AA/L) and polarizable (AMOEBAPRO13 and AMOEBABIO18) force fields have been used. For each one of them, PMFs have been reconstructed and compared in terms of free energy barrier and hydrogen bonding fluctuations behavior over time. Using a SMD velocity of 0.01 Å/ps applied to a set of 20 trajectories, we show that polarizable and non-polarizable force fields do not always agree. As it could be anticipated, strong discrepancies are noticed between polarizable and non-polarizable models when considered in vacuum, whereas results are more comparable when a water environment is added. However, for the largest system, i.e., the CD2CD58 complex, strong differences related to the modeling of a salt bridge are noticed exhibiting some potential issues with classical FFs. Overall, such simulations highlight the importance of the inclusion of polarization effects as PMF free energy barriers computed with AMOEBA always decrease compared to non-polarizable force fields.

INTRODUCTION

Biomolecular systems such as proteins, lipids, or membranes play a considerable role in life processes. It is therefore of a great

Received: February 26, 2019
Published: May 6, 2019

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

importance to be able to compute the free energy associated with the chemical mechanisms occurring for each biological function. In practice, the free energy required by a system to pass from an initial state A to a final state B is strongly dependent on the choice of the reaction coordinate. For example, in the case of docking interaction studies involving ligand protein interactions,[1−3] one can directly envision this reaction coordinate as simply the distance between the two complexes. However, the reaction coordinate is often more complex.[4,5] In all these examples, the free energy profile can be described with the Potential of Mean Force (PMF),[6] which expresses the free energy landscape of the mechanism as a function of the evolution of the reaction coordinate. It is widely used to scrutinize the feasibility of a reaction process.[5−11]

To reconstruct an acceptable PMF for a given biomolecular system, Molecular Dynamics (MD) is a popular tool, but it requires the use of advanced simulation packages[12−17] coupled to a well-chosen sampling strategy. Among the different techniques developed over the years, the Umbrella Sampling (US) approach[12] developed by Torrie and Valleau in 1977 and reviewed in 2011 by Kästner[18] is based on the decomposition of the reaction coordinate in several windows. Each window is restrained by a harmonic potential at certain coordinates owning to the reaction coordinate. The histograms, obtained from each window, can be used to reconstruct the PMF with the help of the Weighted Histograms Analysis Method (WHAM).[19] US and WHAM can be used both with simple reaction coordinates such as distances and with more complex ones such as angles or dihedrals. However, the use of US is not always straightforward and presents some drawbacks as the reaction coordinate is nonuniformly sampled according to the Boltzmann weight and the analysis of the results implies solving nonlinear equations in the WHAM. Furthermore, it is known that a successful application of US depends on the optimal choice of the biasing potential used to retrain the reaction coordinate.

In this context, another methodology has been developed to be employed when the reaction coordinate is limited to a single direction during the biological process: it is called Steered Molecular Dynamic (SMD). Pioneered 20 years ago by the Klaus Schulten's group, it consists of applying an external steering force on one or several atoms to accelerate the desired biological process.[20,21] This steering force is modeled as a harmonic potential restraint. Numerous biological processes have been studied using this method. For example, Izrailev et al. used SMD to analyze and explain the extreme stability of the avidin−biotin complex.[22] They reported a variety of unbinding pathways, especially due to the role of key residues contributing to adhesion as well as the spatial range over which avidin binds biotin. Another example can be found in the Gao et al.'s study which was dedicated to the unfolding of Titin immunoglobulin domains.[23] They demonstrated that the mechanical stability of the domain was linked to the presence of hydrogen bonds within beta strands and to the disulfide bond interactions during the unfolding process. In the same spirit, Bayas et al. also showed that the choice of SMD's setup parameters drives the unbinding pathway of the protein adhesion complex.[24] In their case, it was shown that the CD2−CD58 complex can exhibit two main different unbinding pathways depending on the choice of the used SMD steering force and velocity.

Indeed, if the force applied during the SMD is calculated and the work performed along the reaction coordinate then reconstructed, one issue subsists as it is known that SMD simulations provide nonequilibrium thermodynamic results. To establish a connection between nonequilibrium (SMD) and equilibrium (PMF) properties, Jarzynski overcame the problem by providing a useful relation known as the Jarzynski's equality:[25,26]

$$e^{(-\beta \Delta F)} = \langle e^{(-\beta W_{A\to B})} \rangle \quad (1)$$

where β is equal to $(k_B T)^{-1}$, ΔF represents the free energy difference between the initial (A) and final (B) states, and $\langle . \rangle$ corresponds to an average value and $W_{0\to 1}$ the pulling work performed along the SMD reaction coordinate. It thus relates the free energy difference to an average of pulling works. Nevertheless, even if this equality is of great help, one main problem subsists: the accurate sampling of the exponential average that appears mandatory since this term is dominated by rare trajectories. Therefore, a low velocity/force is required to correctly sample the reaction coordinate. This has the direct consequence of limiting the use of the SMD to only slow processes for which the fluctuation of works is comparable to the temperature.

Several efforts have been attempted to overcome such an issue. One of the most important studies in this direction was performed by Park and Schulten.[27,28] They demonstrated that it was possible to obtain a PMF using a limited set of fast pulling SMD simulations by using the stiff spring approximation methodology coupled to the cumulant expansion approach.[25,27−30] Indeed, it provides an important and non-negligible gain of time in terms of simulation. This method was also used in another work of Yu et al., where the authors employed an SMD set of irreversible simulations to reconstruct the PMF for an amphiphile peptide self-assembly into cylindrical nanofibers.[9] Other works dedicated to SMD technology but in different directions, also exist. For example, Ozer et al. have developed an adaptive mathematical scheme applied to the SMD[31] aiming to generalize the use of Jarzynski's equality through a series of stages that allow for better convergence along nonlinear and long-distance pathways. Thus, a PMF with narrower error bars could be obtained. This methodology has been tested on the unfolding process of neuropeptide Y.[31] Due to the fact that the SMD methodology is strongly direction dependent, being based on the minimization of stretching force, Gu et al. developed a self-adaptive SMD method.[32] It enables finding the best unbinding pathway through the use of a multipopulation genetic algorithm. Overestimations of steering forces are mostly avoided while the reaction coordinate is checked to be the one providing the lowest free energy barrier. It was applied on protein−ligand complexes to reveal the binding affinity without too much overestimation of the stretching force. More recently, Ngo et al. have developed a methodology both based on the Minh−Adib method and the Weighted Histogram Analysis Method.[5] The Minh−Adib method weights forward and backward pulling trajectories in accordance with Crooks' theorem.[33,34] Consequently, it significantly reduces the bias obtained in free energy calculations from nonequilibrium simulations by using an optimal probability distribution on the forward and backward pulling trajectories. Coupled to the nonlinear WHAM equations, it was observed that fluctuations are avoided when fast pulling processes are considered. Such a strategy, applied on an antibiotic gramicidin-A (gA) channel system which is considered as a significant challenge for nonequilibrium sampling, demonstrated that a fast pulling speed (0.01 Å/ps) can return a fair estimate of the PMF of a single potassium ion in gA.[5]

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

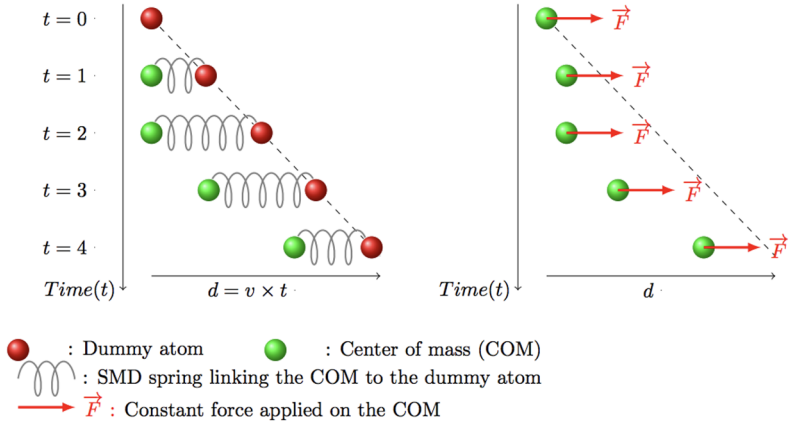


Figure 1. Principle of the constant-velocity SMD (left) and of the constant-force SMD (right). For CVSMD, the SMD atom(s) is (are) in green while the dummy atom is in red, both being linked by a spring (gray). The dummy atom has a linear motion as a function of the time *t*, the harmonic spring enforcing the atoms to follow it. For CFSMD, the SMD atom(s) is (are) in green while a constant force (red arrow) is applied in the same direction at each time-step.

Until now, the SMD method is only implemented in molecular dynamics packages containing non-polarizable force fields[16,35,36] and Drude polarizable models.[37,38] However, despite these efforts, SMD cannot be currently performed with a distributed multipoles/polarizable point dipole force field[39−41] such as AMOEBA,[42−44] while it was demonstrated that directly treating polarization effects provides significant benefits on the thermodynamic properties or parameters' transferability of the polarizable force field. A large number of benchmark studies have been already performed to compare force fields for proteins and lipids.[45−54] In most cases, discrepancies could be observed between the simulations and the experimental data. Therefore, the force field is chosen according to the nature of the biological system and the desired observable, which is a limitation.

The main effort of this present work consists of adapting the SMD methodology to the Tinker-HP massively parallel package[17] in order to offer the possibility of using SMD on large complex systems with advanced point dipole polarizable force fields such as AMOEBA. The paper is organized as follows. First, some methodological aspects concerning the SMD principle and PMFs obtained by SMD simulation at slow (reversible) and fast (irreversible) velocity regimes will be addressed. The use of the stiff spring approximation and the cumulant expansion method described below and developed by Park have been chosen to reconstruct PMF from the irreversible SMD regime. Then we will detail the SMD's implementation inside Tinker-HP and present the systems used for the benchmark procedure, namely, the deca-alanine (104 atoms), the methylated−acetylated deca-alanine (112 atoms), the ubiquitin (9737 atoms), and the CD2−CD58 (97594 atoms) structures. Finally, we will discuss the validity of our SMD implementation using the reversible SMD regime (0.0001 Å/ps) on the alpha-helix stretching of the deca-alanine. Our predicted PMF will be compared to reference values from the literature. Then, following the stiff spring approximation and the cumulant expansion method, PMFs will be reconstructed with a set of 20 trajectories using a fast SMD velocity regime (0.01 Å/ps). This procedure of PMF reconstruction will be assessed using different types of non-polarizable and polarizable force fields for the other benchmark systems. The unfolding of secondary structures such as alpha-helix (deca-alanine and methylated−acetylated deca-alanine) and beta-strands (ubiquitin) as well as quaternary structures (such as ionic salt bridges between two proteins (CD2−CD58)) have been chosen as reaction coordinates for the SMD simulations providing incremental studies in terms of size and systems complexity. It also allows us to compare the behavior of hydrogen bonds, ionic interactions, and the stability of secondary and quaternary structures within each class of non-polarizable and polarizable force fields.

METHODOLOGY

In this section, the theoretical framework of the Steered Molecular Dynamics (SMD) is described. In particular, we detail how to obtain a Potential of Mean Force (PMF) based on SMD calculations. A first strategy uses a fully reversible regime and requires a small SMD velocity (0.0001 Å/ps) whereas a second possibility uses faster SMD velocities in a non-equilibrium regime resorting to the Jarzynski's equality. The latter can be coupled to several approximations. Finally, the implementation and the performance of this method in the Tinker-HP software are presented.

Link between SMD and PMF. Free energy simulation of biochemical systems has for years been an intense field of research.[55,56] On the one hand, to perform such computations, one can choose to use equilibrium methods such as Umbrella Sampling (US)[12] or Thermodynamic Integration.[57] On the other hand, the SMD method, pioneered 20 years ago by the Schulten's group,[27,28] allows making a bridge between non-equilibrium simulations (SMD) and equilibrium properties (PMF). It consists of applying an external force on the system by employing only a harmonic spring and a dummy atom. As depicted on Figure 1, one or several atom(s) (in green here) are linked to the dummy atom (in red) by a harmonic spring (in blue), named as the SMD atom(s). The dummy atom moves at constant velocity as a function of time, and the harmonic spring enforces the atoms to follow it. Therefore, assuming that in an MD simulation covalent bonds could not be broken, all the atoms directly linked to the SMD atom(s) will follow the motion. This method is called constant-velocity SMD (CVSMD). Another related method also exists, which induces a constant force on the SMD atom(s) (in green). This method is called constant-force SMD (CFSMD) and is illustrated in Figure 1. Let us consider a classical system of N particles in a canonical ensemble in contact with a heat bath of constant temperature T. Such system is characterized by $3N$ positions r and momenta p.

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

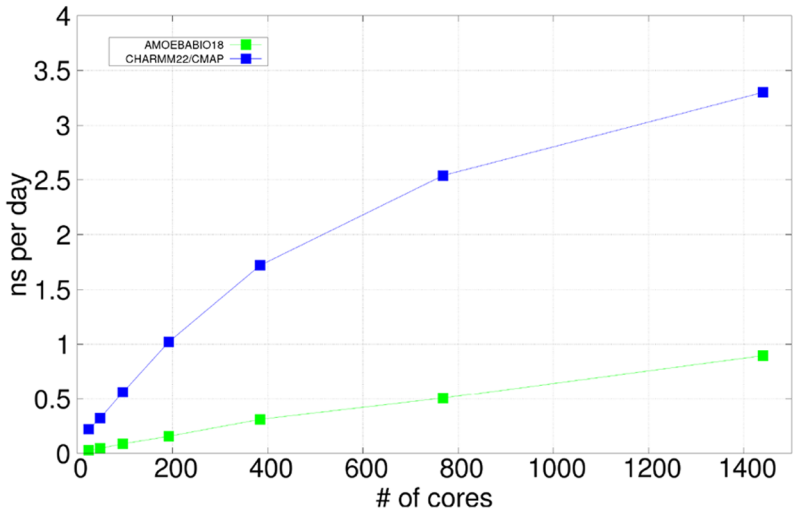


Figure 2. Performance gain for the SMD stretching of the CD2−CD58 complex in water (97594 atoms) using the non-polarizable force field CHARMM22/CMAP and the polarizable AMOEBABIO18 force field.

The SMD procedure adds a harmonic potential energy term to the original Hamiltonian energy. The Hamiltonian of the SMD system $H'(r, p, \lambda)$ can be written as a sum of two contributions:

$$H'(r, p, \lambda(v, t), t) = H(r, p) + h(r, \lambda(v, t)) \quad (2)$$

with $H(r, p)$ the Hamiltonian of the system at equilibrium and $h(r, \lambda(v, t))$ the SMD harmonic potential energy defined as

$$h(r, \lambda(v, t), t) = \frac{k}{2}(\epsilon(r) - \lambda(v, t))^2 \quad (3)$$

with k represents the spring constant, $\epsilon(r)$ the reaction coordinate related to the steered atom position r at time t, and $\lambda(v, t)$ an external parameter corresponding to the displacement of the dummy atom; $\lambda(v, t) = \lambda_0 + v \times t$ with v the velocity of the dummy atom and t the time of the simulation. The work done on the system as a function of time is then

$$W_{0\to t} = -k \times v \int_0^t \mathrm{d}t'(\epsilon(r) - \lambda_0 - v \times t) \quad (4)$$

This work $W_{0\to t}$ can be considered as the relative free energy ΔF between the initial and final states only when the SMD velocity is small enough to consider the process as reversible. Otherwise, for faster ones, according to the second law of thermodynamics, this work overestimates the final relative free energy ΔF.

From Irreversible to Reversible Processes. Thus, being in a reversible regime lead to an important computational cost, especially with large biological systems. The SMD velocity is then chosen to be higher, and in practice, the regime passes from reversible to irreversible. The position of the dummy atom $\lambda(v, t)$ fluctuates around the reaction coordinate $\epsilon(r)$, but when one uses a sufficiently large spring constant k for the guiding potential, Park et al.[27] showed that the free energy of the process can be written directly as a function of the PMF, labeled as $\Phi(\epsilon)$. When $\epsilon = \lambda$, one has then $F(\lambda) = \Phi(\lambda)$. This formula is known in the literature as the stiff spring approximation.[27,28] Using the Jarzynski equality, one has then:

$$\Phi(\lambda_0) - \Phi(\lambda_t) = \frac{1}{\beta} \log\langle e^{(-\beta W_{0\to t})}\rangle \quad (5)$$

The exponential average being dominated by rare trajectories, the application of the cumulant expansion method (CE) has been proven to be efficient when the work is of "gaussian" nature.[25,27,29] Applied on eq 5, one has the formula:

$$\Phi(\lambda_t) = \Phi(\lambda_0) + \langle W\rangle - \frac{\beta}{2}(\langle W^2\rangle - \langle W\rangle^2) + \frac{\beta^2}{6}(\langle W^3\rangle - 3\langle W^2\rangle\langle W\rangle + \langle W\rangle^3) + \ldots \quad (6)$$

As explained before by Park et al.,[27,28] using the CE provides two types of errors: a systematic error due to the truncation of the CE order and a statistical error due to the sampling.

Implementation in Tinker-HP and Available Force Fields. To be able to treat efficiently large biological systems, software allowing for SMD has to be massively parallel. As we pointed out previously, it is also interesting to be able to use SMD with second generation polarizable force fields such as AMOEBA, which have proven their efficiency in terms of accuracy and transferability. This is why we implemented the SMD in Tinker-HP,[17] the massively parallel MD engine present within the Tinker package (now version 8[58]). For performance purposes, the SMD has been directly implemented within the 3D domain decomposition parallel framework of Tinker-HP to optimize communications and therefore performances to ensure that the scaling of a SMD calculation stays similar to that of a regular MD simulation.[17] The SMD methodology was also implemented in the canonical, non-MPI Tinker code,[58] and a comparative test on deca-alanine and ubiquitin was performed on both pieces of software to ensure that the domain decomposition implementation correctly reproduces forces in Tinker-HP. Results were found to be fully identical. Various classical force fields such as AMBER99, AMBER99SB, CHARMM22/CMAP, and OPLS−AA/L have been tested. The classical water model used in all simulations is TIP3P.[59] Two AMOEBA protein force field parametrizations are also assessed, the initial AMOEBAPRO13[43] and the recent (reference) reparametrization AMOEBABIO18.[43,44] In all cases, we used the 2003 AMOEBA water model.[42] The SMD implementation will be made available free of charge for academics on the Tinker-HP Web site (http://tinker-hp.ip2ct.upmc.fr/) and later on github (https://github.com/TinkerTools/Tinker-HP).

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

RESULTS AND DISCUSSION

SMD's Scaling within Tinker-HP. The SMD implementation and scalability within the Tinker-HP v1.1 software have been tested on a reasonably large system, i.e., the CD2−CD58 complex encompassing 97594 atoms (including water molecules) (Figure 2). Tests used the Velocity Verlet integrator, the Bussi thermostat, and a time step of 1 fs with a polarizable force field (AMOEBABIO18) and a non-polarizable force field (CHARMM22/CMAP). For AMOEBA, the polarization contribution was computed using a conjugate gradient solver (with a diagonal preconditioner)[60,61] but without resorting to a predictor corrector to avoid any dipole history interferences during our SMD testings. The water molecules were kept fully flexible, and no restraint was employed on the hydrogens beared by heavy atoms (no rattle/shake). Smooth Particle Mesh Ewald (SPME) cutoffs have been set to 7 Å for the real space contribution, the van der Waals cutoff being set to 9 Å. SPME grids were chosen to be 96 Å × 96 Å × 192 Å. The purpose of these scaling curves is not to measure the best performances of Tinker-HP. Indeed, we did not intend here to use the available more aggressive setups especially for the AMOEBA part[17,62] as all SMD tests have used very conservative settings for both classical and polarizable benchmarks. Also, one has to note that the non-polarizable simulations were performed on Tinker-HP 1.1 and are not as optimized as the polarizable ones. Despite the absence of activation of several available acceleration techniques, the observed simulation timings are reasonable for such a system, with a performance of 0.9 and 3.3 ns per day when 1440 cores are used respectively with AMOEBABIO18 and CHARMM22/CMAP. These results demonstrate the possibility to perform SMD simulations on large systems using polarizable and non-polarizable force fields such as AMOEBA and CHARMM. We also confirm that the SMD constraints have no impact on the code scalability and performances.

First Benchmark: The Gas-Phase Deca-alanine. We first benchmarked our SMD implementation on the gas-phase deca-alanine (Figure 3) system, a small oligopeptide made of 10 alanine, encompassing a total of 104 atoms, and being characterized by only one alpha-helix. The structure has been chosen to be exactly the same as the one Park et al. used in their reference work.[27] The N terminal atom of the first residue localized at the origin, and the C terminal of the 10th residue aligned along the z axis. The distance between these atoms is the reaction coordinate. The alpha-helix's stretching of the deca-alanine has been performed with the CVSMD methodology with two different velocity regimes:

- a reversible regime with an SMD velocity of 0.0001 Å/ps applied during a single simulation of 200 ns. The SMD simulations have been performed with different integrators (Velocity Verlet,[63] RESPA,[64] BAOAB[65]) and two different time steps, i.e., 1 and 2 fs, in order to ensure the validity of the SMD in all these different cases. We used a Bussi thermostat.[66] All these results have been compared to the ones obtained with the reference NAMD implementation of SMD.[16]

- an irreversible regime with an SMD velocity of 0.01 Å/ps and requiring the evaluation of a set of different trajectories. Each independent trajectory is about 2 ns and generated with a different initial condition in the canonical ensemble. In this case, we used the stiff spring approximation and CE method.

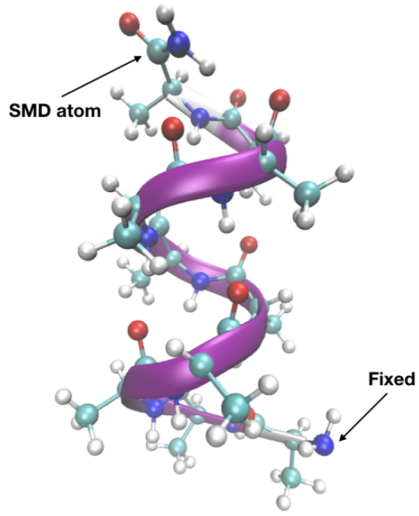


Figure 3. Representation of the deca-alanine system (atoms are depicted with the CPK representation while the alpha-helix is depicted as a purple newribbon representation; structure adapted with permission from Park et al.[27] Copyright 2003 American Institute of Physics). During the CVSMD simulations, the N terminal atom of the first residue is kept fixed whereas the C terminal atom of the 10th residue is pulled in the z direction with a constant velocity.

The CHARMM22/CMAP force field was chosen to compare our results to those extracted from the work of Park et al.[27,28] The N terminal atom of the first residue has been kept fixed, while the C terminal atom of the 10th residue has been pulled in the z direction. A spring constant of 7.2 kcal/(mol·Å^2), coupled to a transverse spring constant with the same value, has been considered. No cutoff was involved for nonbonded interactions.

Finally, PMFs generated from the reversible and irreversible regimes have been compared to the one obtained by using Umbrella Sampling (US), which can be considered as a reference.

PMFs from Reversible Pulling. To validate the SMD implementation in the Tinker-HP software, we chose to induce the unfolding of the deca-alanine oligopeptide in a reversible manner. It requires a 200 ns simulation for the full extension for each unfolding simulations. Several time steps (1 and 2 fs) and integrators (RESPA, Velocity Verlet, and BAOAB) have been tested with Tinker-HP and compared with the NAMD's SMD with a RESPA integrator and time steps of 1 and 2 fs. As depicted in Figure 4, all the Tinker-HP PMFs obtained for a time step of 1 fs appear in agreement with the NAMD PMFs of 1 fs. When a time step of 2 fs (RESPA) is used in Tinker-HP, some minor differences appear after the end-to-end distance of 25 Å, with a difference of 0.5 kcal/mol compared to the 1 fs PMFs. This little energy gap is also observed in the case of the NAMD RESPA of 2 fs, showing that the choice of the time step has some limited impact in terms of long time scale free energy calculation. Nevertheless, the Tinker-HP BAOAB PMFs with a time step of 2 fs are in good agreement with the ones obtained with a time step of 1 fs.

To further investigate the previously observed small energetic gap and to ensure that it is not a consequence of a too high SMD velocity choice (i.e., nonequilibrium regime), the reverse process was also performed. It corresponds to the deca-alanine folding (or unstretching), the starting structure being the last one obtained by the unfolding simulation. The same simulation parameters were used. Figure 5 shows that the PMFs obtained

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

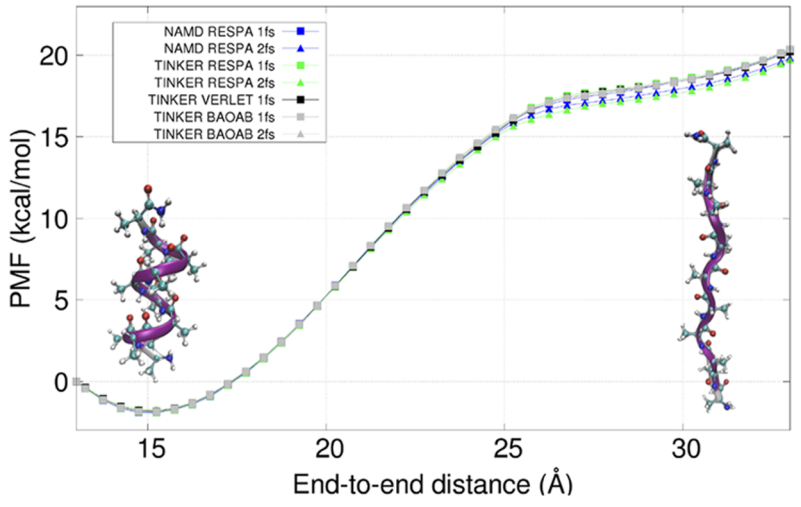


Figure 4. Tinker-HP and NAMD results for the CVSMD on the deca-alanine with CHARMM22CMAP. Several timesteps (1 and 2 fs) and integrators (RESPA, Velocity Verlet, and BAOAB) have been tested. A movie depicting such simulations is available in the Supporting Information (SI-1).

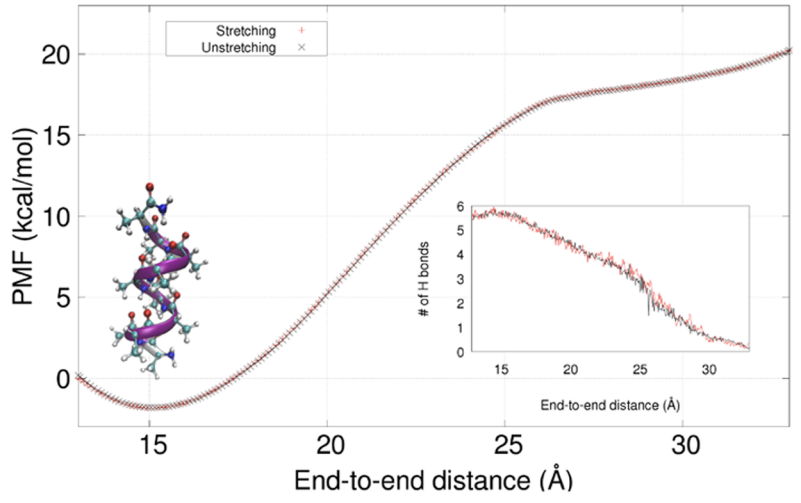


Figure 5. Stretching (red) and unstretching (green) process of the gas-phase deca-alanine with a pulling rate of 0.0001 Å/ps. Hydrogen bonds (HBs) behavior during this process is depicted and has been calculated with a cutoff radius of 3.5 Å and an angle cutoff of 40°.

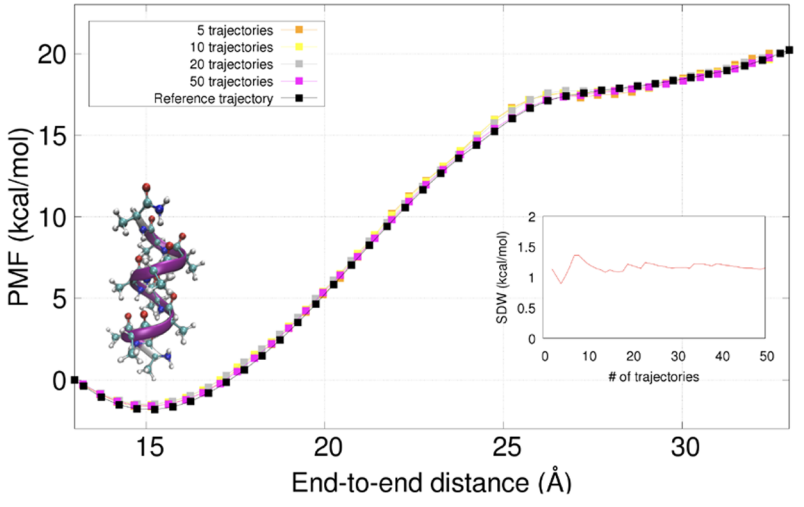


Figure 6. PMFs obtained for different sets of trajectories (5, 10, 20, 50) with Tinker-HP on the deca-alanine system. The black curve corresponds to the reversible PMF. Evolution of the standard deviation work (SDW) is also depicted as a function of the number of trajectories on the bottom right.

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

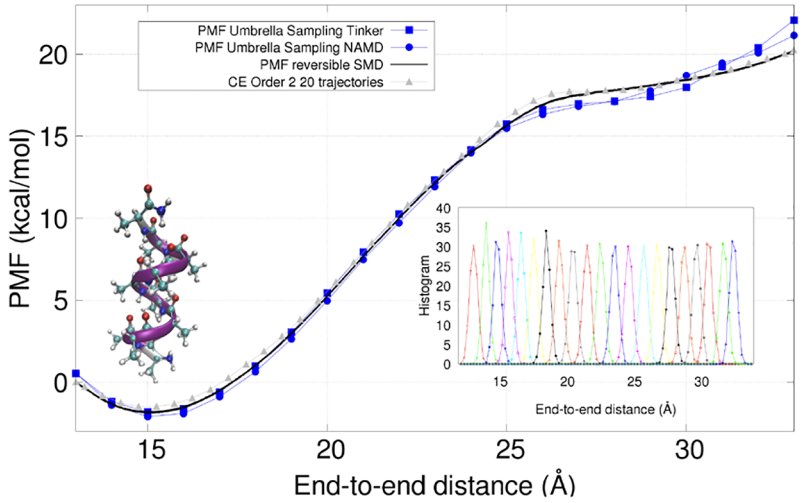


Figure 7. Direct comparison for the stretching process of the deca-alanine: reversible SMD simulation (black line) and PMF obtained with a set 20 trajectories in an irreversible regime vs Umbrella Sampling (blue lines).

during the unfolding and folding simulation have a very good overlap. Since the same PMF could be obtained from a forward pulling (unfolding case here) and a reverse pulling (folding case here), the SMD simulations can be considered as reversible for this system. The HBs' behavior, also depicted in Figure 5, supports this conclusion because their overlap indicates a similar behavior of the alpha-helix during the unfolding and folding simulations. All these results are in full accordance with the work of Park et al.[27] and their conclusions on the reversibility feature of the SMD velocity (0.0001 Å/ps) on the deca-alanine system.

PMFs from Irreversible Pulling. The reversible regime is easily reached for small systems. Nevertheless, when studying large systems such as biomolecules, simulation time becomes a crucial issue. In this case, a single 200 ns simulation is possible but not computationally easily accessible. However, instead of resorting to only one reversible simulation, a set of several irreversible simulations can be performed and the PMF finally reconstructed using the stiff spring approximation and the cumulant expansion at the second order. We changed the SMD velocity from 0.0001 Å/ps to 0.01 Å/ps, and so the simulation time went from from 200 to 2 ns per trajectory giving access to an important gain in terms of simulation time. A total number of 20 and then 50 independent pulling simulations have been performed corresponding to a total of 40 ns of regular Molecular Dynamics simulation time. Although the number of 20 trajectories was empirically chosen by us, different sets of trajectories (5, 10, 20, 50) have been considered in order to estimate the best number of trajectories required to ensure an accurate estimation of the PMF while limiting the statistical error. The standard deviation work (SDW) $\sqrt{\langle W_{\text{total}}{}^2\rangle - \langle W_{\text{total}}\rangle^2}$, where $\langle W_{total}{}^2\rangle$ the average of the squared of the total work and $\langle W_{total}\rangle^2$ the squared average of the total work, has been used as a direct qualitative measure of the convergence of the PMF with the number of trajectories. As depicted on Figure 6, only 5−10 trajectories seem to be enough to obtain a PMF similar to our reference. However, our choice of 20 trajectories is supported by the behavior of the SDW, where a convergence to 1.2 kcal/mol SDW begins to be observed when 20−30 trajectories are considered. This result is in accordance with the standard deviation value found by Park et al., with a value of 1.9 kcal/mol in the same conditions but by considering only 10 instead of 20 trajectories. The PMF has been evaluated to be equal to 21.5 kcal/mol in the reversible regime, and the standard deviation value thus corresponds to 11.6% of the PMF value.

Umbrella Sampling Comparison. To obtain PMFs, several methodologies other than SMD exist. For example, despite being more computationally intensive, Umbrella sampling is a popular method to obtain an accurate reference PMF,[12] and it is thus of interest to compare our SMD results to it. Twenty windows of the distance reaction coordinate have been generated, in combination with a harmonic potential. The constant spring used to restrain the system is the same as the one used before in the SMD studies, and the chosen biased end-to-end distance for the windows varies from 13 to 33 Å, with a step of 1 Å between each biased end-to-end distance. Each trajectory within a window is 2 ns long, corresponding to a total of 40 ns of simulation. Histograms were obtained and combined with the Weighted Histogram Analysis Method (WHAM) implemented by Grossfield[67] to generate the PMF and to compare it with the reversible and irreversible SMD results obtained previously (Figure 7). The curves are in good agreement, similarly to what was pointed out by Hummer and Park.[27,28,30] However, a little discrepancy can be observed at the end-to-end distance of 30 Å where the US PMF is slightly above both SMD ones. This was not as clearly reported in Park's results[27,28] but was fully discussed in a study by Hazel.[68] To conclude on this issue, the US PMF was calculated with NAMD using the same procedure and parameters used in Tinker. The results depicted in Figure 7 show that the same behavior is found for both US curves using the different pieces of software validating the Tinker-HP implementation. This discrepancy being only observed at high values of the reaction coordinate when the alpha-helix is completely unfolded, it can be considered as nonsignificant. At this point and after this series of initial tests leading to results in agreement with the literature, we can draw conclusions about the validity of our new SMD implementation in Tinker-HP.

SMD: Polarizable vs Non-Polarizable Force Field Comparisons. *Methylated/Acetylated (M−A) Deca-alanine.* The second chosen system is also a deca-alanine, but with two differences on top of the residues 1 and 10, where the NH_2 and $CONH_2$ end functions are replaced by an acetylated and a methylated end function. The final system is made of 112 atoms

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

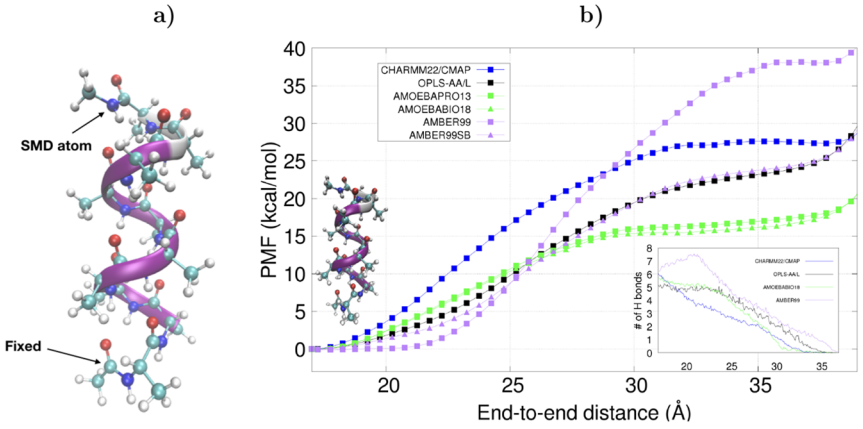


Figure 8. (a) Representation of the methylated/acetylated (M−A) deca-alanine. (b) Force field comparison for the stretching process of the methylated/acetylated (M−A) deca-alanine. PMFs have been obtained with a set of 20 trajectories with an SMD velocity of 0.01 Å/ps (irreversible regime). The average number of hydrogen bonds (HBs) for CHARMM22/CMAP, AMBER99, OPLS−AA/L, and AMOEBABIO18 force fields are also depicted on the bottom right and have been calculated with a radius cutoff of 3.5 Å for the donor/acceptor distance and an angle cutoff of 40°.

and still considered within the gas phase. This change is justified by the usual parametrization of these end caps in most force fields, while the previous non-usual NH_2 and $CONH_2$ end functions were only parametrized in CHARMM22/CMAP by Park et al. for their work.[27] Initial structures were generated with the TINKER PROTEIN program, with $\Phi = -60°$ and $\Psi = -43°$ for torsion angles, corresponding to the standard right-handed alpha-helix. Ten ALA (alanine) groups have been added, with ACE and NME capping groups at each end. This procedure was applied for every force field considered, AMBER99,[69−73] AMBER99SB,[69−74] CHARMM22/CMAP,[75−77] OPLS−AA/L,[78,79] AMOEBAPRO13,[43] and AMOEBABIO18.[43,44] We applied the same protocol previously performed for the deca-alanine for the alignment, minimization, and equilibration of the M−A deca-alanine. Finally, 20 independent SMD trajectories were performed with an SMD velocity of 0.01 Å/ps and a spring constant of 7.2 kcal/(mol·Å^2), coupled to a transverse spring constant with the same value. The C carbonyl carbon of the first residue was kept fixed where the N amide nitrogen of the 10th residue was pulled in the z direction. PMFs have been thus obtained by using of the stiff spring approximation and the cumulant expansion at the second order. Knowing that there exists differences in the literature about the secondary structure's description and behavior of proteins (principally alpha-helix and beta strands) as a function of the chosen force field,[49−54,80] it is of interest to calculate the PMF of the current single alpha-helix with a set of different generation force fields (polarizable and non-polarizable).

As represented on Figure 8, each force field is associated with a specific PMF and to a specific free energy barrier. Compared to the initial deca-alanine, this M−A deca-alanine depicts a little higher free energy barrier. For example, it is around 28 kcal/mol with the CHARMM22/CMAP force field (with a SDW of 1.40 kcal/mol, see Table 1) while the previous free energy barrier for the deca-alanine[27,28,68] was found to be equal to 21.5 kcal/mol. The only differences being located on the end-caps used with the M−A deca-alanine and in the choice of the SMD and fixed atoms, this sizable difference in the free energy barrier of the deca-alanine and M−A deca-alanine with CHARMM22/CMAP can be attributed to the difference in reaction coordinates. Indeed, it appears that it is defined at an oblique angle to the helix axis (see Figure 3), yet the M−A deca-alanine has a reaction coordinate coincident with the helix axis due to the choice of atoms (see Figure 8). Still, it appears that the number of hydrogen bonds is similar between the two types of structures at their equilibrium state (see Figure 5 and Figure 8).

Table 1. Free Energy Barrier and Standard Deviation Work (SDW) Calculated with a Set of 20 Trajectories for Each Force Field for the M−A Deca-alanine[a]

force field	free energy barrier (in kcal/mol)	SDW (in kcal/mol)
AMBER99	38.0	1.76
AMBER99SB	24.8	0.65
OPLS−AA/L	24.5	0.82
CHARMM22/CMAP	27.4	1.40
AMOEBAPRO13	18.4	1.09
AMOEBABIO18	17.1	1.15

[a]The free energy barriers correspond to the free energy obtained when the plateau is reached (so when the alpha-helix is completely unfolded).

Interestingly, PMFs obtained for all the force fields considered here show a different behavior, ranging from 18 to 38 kcal/mol in terms of free energy barrier (see Table 1). Even if the results obtained with the two AMOEBA force fields (AMOEBAPRO13 and AMOEBABIO18) are a little different, they fall in the same range of final free energy barrier of approximately 18−19 kcal/mol (with an SDW encompassed between 1.10−1.20 kcal/mol, see Table 1) which is the lowest compared to the other force fields. The relatively small SDW values show that 20 trajectories seem to be a good choice for this case. AMBER99SB and OPLS−AA/L provides both a free energy barrier of 25 kcal/mol with the lowest SDW values, but the most tricky case is given by AMBER99, which depicts an astonishing behavior compared to the other tested force fields. Indeed, while the number of HBs should decrease during the stretching process, it increases for AMBER99. This additional HB was found to occur between the NH amide of the NME capping group and the CO carbonyl of the 8th residue during the pulling process. To confirm these results and ensure that this HB is not due to a preparation artifact, we repeated the set up procedure 10 times for this force field and found 10 times the same initial structure as a starting point, still without this HB. Such behavior can be associated with a force field defect. Its high free energy barrier (38 kcal/mol) and

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

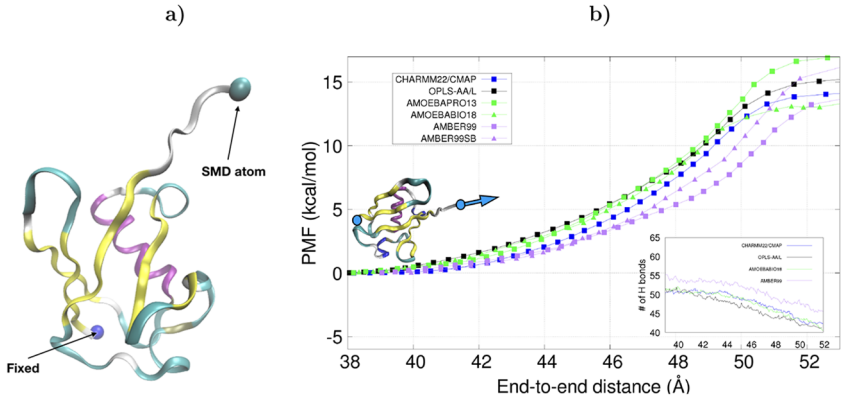


Figure 9. (a) Representation of the ubiquitin system (three beta strands (in yellow) and the alpha-helix (in purple); structure coming from the 1UBQ protein data bank file). During the CVSMD simulations, the N terminal atom of the first residue is kept fixed whereas the C terminal atom of the last residue is pulled in the direction formed by the vector crossing both atoms with a constant velocity. (b) Force field comparison for the stretching process of the beta strand of the ubiquitin. All the PMFs have been obtained with a set of 20 trajectories with an SMD velocity of 0.01 Å/ps (irreversible regime). The average HBs behavior is also depicted at the bottom right and has been calculated with a radius cutoff of 3.5 Å and angle cutoff of 40°. A movie depicting such simulations is available in the Supporting Information (SI-2).

especially its higher SDW (1.76 kcal/mol) suggests that this PMF overestimates the alpha-helix stability of the M−A deca-alanine in the vacuum and also that 20 trajectories are not sufficient for this case. Thus, decreasing the SMD velocity or increasing the number of trajectories are two possible ways to get a more accurate result for this case.

Nevertheless, having a free energy barrier with a low SDW is not a guarantee that the value is converged. However, it clearly appears that stability of alpha-helix in the vacuum is very different between the type of chosen force field. The choice of force field to describe the secondary structure of proteins in vacuum thus seems to be crucial, but this could also suggest that these force fields are not suitable to the gas phase and should be only used in a solvated environment. Of course, such a conclusion is known as all classical force fields are designed for the condensed-phase "only" whereas polarizable force fields can "adapt" to different conditions thanks to the inclusion of the polarization contribution.

Ubiquitin. Made of one alpha-helix and of three beta-strands, the ubiquitin protein is a more complex and realistic system that we studied in the condensed phase, i.e., considering a solvation water box. In this case, the reaction coordinate chosen for SMD simulations is the distance between the C terminal atom of residue 76 (GLY) and the N terminal atom of residue 1 (MET), corresponding to the stretching of one of its beta-strands (see Figure 9). The same force fields used before have been considered. The system has been solvated in a periodic waterbox of dimension 55 Å × 42 Å × 42 Å and is thus made of a total of 9737 atoms, 1232 belonging to the protein structure. In terms of complexity, the system size is scaled by a factor 100 compared to the previous system, and the role of water molecules is directly included in the final PMF. The Smooth Particle Mesh Ewald (SPME) method was used, with a real space cutoff of 7 Å and a SPME grid of 60 Å × 45 Å × 45 Å. VDW interactions were used with a cutoff of 9 Å. The spring constant has been chosen to be equal to 7.00 kcal/(mol·Å^2). A second spring, with the same constant as the first one, has been included. The SMD velocity was equal to 0.01 Å/ps and applied on the C terminal of residue 76 (GLY), when the N terminal of residue 1 (MET) was kept fixed (fixed atom). The SMD pulling direction has been chosen to be along the vector formed by the fixed and the SMD atom. Twenty independent SMD trajectories have been generated for each force field, and PMFs have been generated with these as for the M−A deca-alanine. According to Table 2, the different calculated free energy barriers are in better agreement compared to the previous system (M−A deca-alanine in vacuum, see Table 1), ranging from 13.2 kcal/mol for the AMOEBABIO18 force field to 16.5 kcal/mol for the AMOEBAPRO13 force field, corresponding to a statistical range of 3.3 kcal/mol (where the M−A deca-alanine provided a critical value of 20.9 kcal/mol (see Table 1)). AMOEBABIO18 should be clearly taken as a reference as lots of work has been performed on the 2018 parametrization to improve the quality of the multipolar electrostatics in order to be consistent with the nucleic acids addition.[44] Such results clearly show the improvements of the recent refinements. The SDWs are close together and consistent, ranging from 0.76 to 1.16 kcal/mol. Representing about 6% of the free energy barriers, it means that the choice of 20 trajectories for this system is satisfactory. It could also be reported that the free energy barriers for the unfolding of ubiquitin are lower than for the unfolding the M−A deca-alanine. The systems being completely different (M−A deca-alanine is a fictive alpha-helix in vacuum, where ubiquitin is a fully solvated protein), direct comparisons of their respective free energy barriers make no sense. With a difference of 3 kcal/mol for the free energy barrier and a negligible variation of the

Table 2. Free Energy Barrier and Standard Deviation Work (SDW) Calculated with a Set of 20 Trajectories for Each Force Field for the Ubiquitin Protein[a]

force field	free energy barrier (in kcal/mol)	SDW (in kcal/mol)
AMBER99	13.3	0.76
AMBER99SB	15.4	0.95
OPLS−AA/L	15.0	1.15
CHARMM22/CMAP	14.0	1.16
AMOEBAPRO13	16.5	0.94
AMOEBABIO18	13.2	0.86

[a]The free energy barriers correspond to the free energy obtained when the plateau is reached (so when the beta-strand is completely unfolded).

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

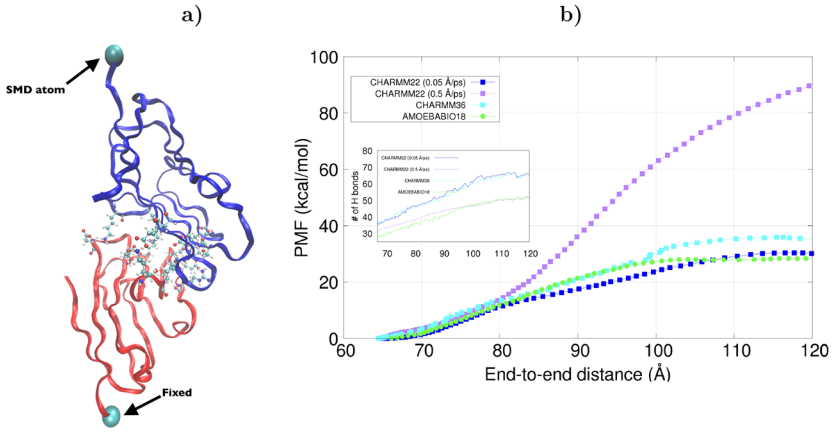


Figure 10. (a) Representation of the CD2−CD58 complex (the CD2 part (in blue) and CD58 (in red) are depicted with the newribbon representation and the salt bridges are depicted with the bond representation; structure coming from the 1QA9 protein data bank file). During the CVSMD simulations, the backbone carbon atom of residue 95 of the CD58 part is kept fixed whereas the backbone C atom of residue 105 of the CD2 part is pulled in the z direction with a constant velocity. (b) PMFs of the salt bridges' breaking process of the CD2−CD58 complex. A set of 20 trajectories has been considered for each PMF. The CHARMM22 non-polarizable force field has been chosen with two different SMD velocities (0.5 and 0.05 Å/ps) while the CHARMM36 non-polarizable and AMOEBABIO18 polarizable force fields have been chosen with only one SMD velocity (0.05 Å/ps). The inset represents the evolution of the hydrogen bonds formed between each amino acid of the salt bridges and the water molecules of the solvent as a function of the end-to-end distance. Movies depicting such simulations are available in the Supporting Information (SI-3, mechanism A, and SI-4, mechanism B).

SDW values for the AMOEBA family of force fields, these results denote a small but non-negligible difference between the AMOEBABIO18 and the AMOEBAPRO13 force fields whereas no differences were observed in the previous case between them. With a lower free energy barrier and SDW, AMOEBABIO18 tends to decrease the free energy barrier, going in the same direction as observed before for the M−A deca-alanine test case. However, AMOEBAPRO13 increases the free energy barrier compared to all the other force fields in this case, where it was not observed for the M−A deca-alanine test case. It thus suggests a better confidence in AMOEBABIO18 for future work. The number of HBs is also represented in the inset of Figure 9 for the CHARMM22/CMAP, OPLS−AA/L, AMOEBABIO18, and AMBER99 force fields. Even if 5 more HBs are initially observed for the AMBER99 case (55 instead of 50), each force field exhibits the same behavior, with a total breaking of 9 hydrogen bonds, corresponding to the total breaking of the beta-strand at the end-to-end distance of 52 Å. We can then state that all considered force fields predict a similar unfolding process of the ubiquitin's beta-strand. The addition of a water environment clearly shows the importance of considering the solvent within SMD comparisons, and we then decide to strongly increase the size of the protein system and therefore the amount of water molecules solvating such systems to access a more realistic system: the CD2−CD58 complex.

CD2−CD58. Salt bridges (labeled as SBs in the following) are, along with alpha-helices and beta-strands, responsible for the protein's stability, structure, and design.[81,82] Their accurate modeling using force fields and the full understanding of their mechanisms remain an important computational challenge.[83−89] To tackle such systems and to increase the complexity of our test cases, we decided to simulate the human CD2−CD58 complex. Characterized by SBs between two complexes, it was already studied with SMD by Bayas et al. with the CHARMM22 non-polarizable force field to scrutinize the molecular details underlying the complex detachment of the SBs.[24] The SMD methodology allows us to get an approximate PMF with these kinds of system, which is not feasible by using other methodologies such as US in terms of computational time, especially when using polarizable force fields. The dual aim of this benchmark is to test the SMD on a much larger system with both polarizable and non-polarizable force fields, while also evaluating the scalability of the implementation in Tinker-HP. The force fields chosen for this study were CHARMM22 and AMOEBABIO18. The CD2−CD58 was first aligned along the z axis, where the backbone carbon of the residue 105 of CD2 was translated to the origin and the z axis aligned with the vector formed by the backbone carbons of the residue 105 and 95 of respectively CD2 and CD58 (see Figure 10). The system was solvated in a water box of 80 Å × 80 Å × 160 Å, corresponding to 97594 atoms, 3223 belonging to the CD2−CD58 complex. The spring constant has been chosen to be equal to 1.00 kcal/(mol·Å^2). A second transverse spring, with the same constant as the first one, has been included. The SMD velocity was equal to 0.05 Å/ps and applied on the backbone carbon of residue 105 of CD2, while the backbone carbon of residue 95 of CD58 was kept fixed. Twenty independent SMD trajectories (20 × 1.3 ns) have been generated using each force field, and PMFs have been generated with these as for the previous systems.

According to Figure 10, the free energy barriers for the CHARMM22 and AMOEBABIO18 force fields with the SMD velocity of 0.05 Å/ps are similar within a range of 1.5 kcal/mol (the figure also depicts results for an SMD velocity of 0.5 Å/ps for CHARMM22 and 0.05 Å/ps for CHARMM36 that are discussed further in this section). However, and strikingly, their respective SDW values are remarkably different, ranging from 2.38 kcal/mol for CHARMM22 to 5.40 kcal/mol for AMOEBABIO18. Although previous SDW values have been shown to be around only 1 kcal/mol, it demonstrates that the empirical choice of 20 trajectories is not enough for this system, and thus more than 20 trajectories needs be considered. The addition of three trajectories for the AMOEBABIO18 case supports this claim, as it makes the SDW decrease, passing from 5.40 to 5.21 kcal/mol. However, due to the high computational

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

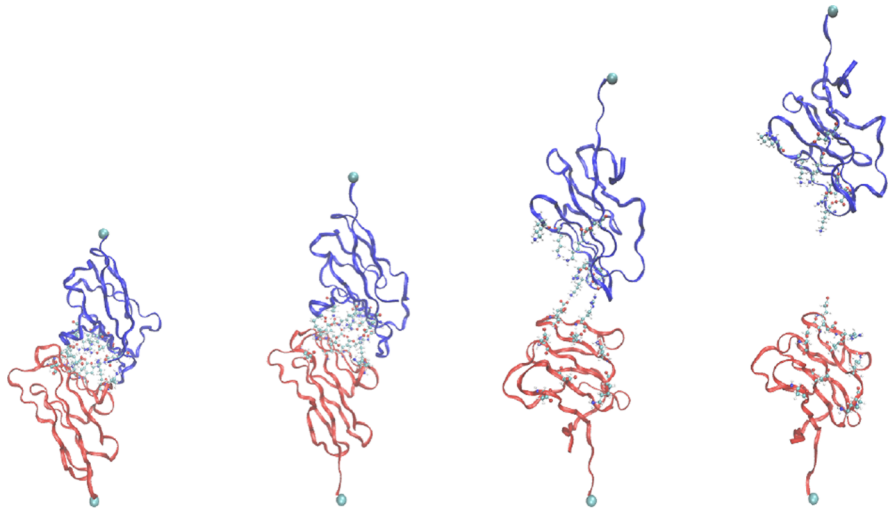


Figure 11. Mechanism A: four snapshots taken at different end-to-end distances (from left to right: 66 Å, 83 Å, 110 Å, and 120 Å) during a SMD simulation performed on the CD2−CD58 complex with the CHARMM22 force field with an SMD velocity of 0.05 Å/ps. The complex dissociated in this case without important modifications of the secondary structures of both subcomplex parts. A movie depicting such simulations is available in the Supporting Information (SI-3).

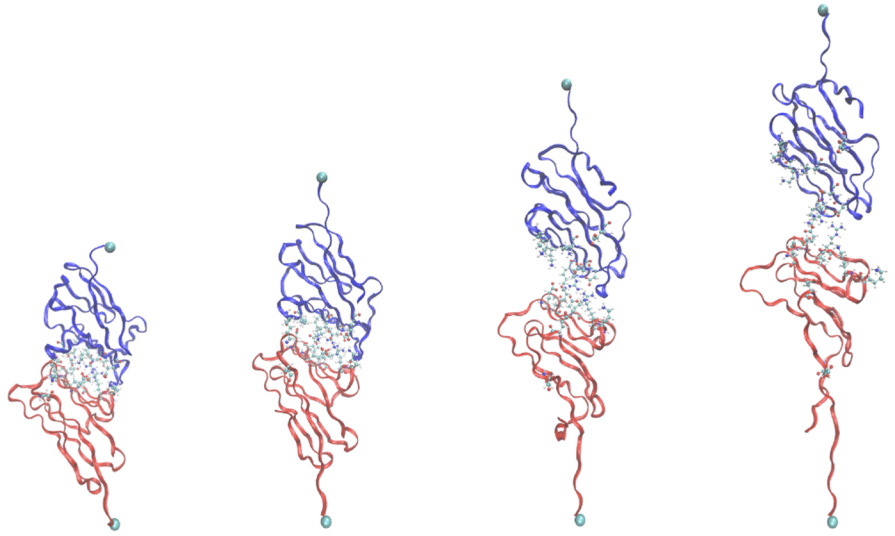


Figure 12. Mechanism B: four snapshots taken at different end-to-end distances (from left to right: 66 Å, 83 Å, 110 Å, and 120 Å) during an SMD simulation performed on the CD2−CD58 complex with the AMOEBABIO18 force field with an SMD velocity of 0.05 Å/ps. The complex does not completely dissociate in this case, showing important modifications of the secondary structures on the CD58 subcomplex parts (in red). A movie depicting such simulations is available in the Supporting Information (SI-4).

cost for only one trajectory for this system with the AMOEBABIO18 force field, it remains a hard task to really increase the number of trajectories.

This high energetic difference for the SDW values could also be related to the complexity of the unfolding pathway. Instead of only one mechanism as in the previous cases, two different mechanisms were observed and are in competition to describe the breaking process of the CD2−CD58 cell adhesion complex:

- Mechanism A: Depicted in Figure 11, CD2−CD58 begins to break its SBs which are localized along the z axis. It is followed by the alignment along the z axis of the other SBs which were originally perpendicular to the z axis. Once they are aligned they totally break without inducing any denaturation of secondary structures of the two subparts of the complex.
- Mechanism B: Depicted on Figure 12, CD2−CD58 begins to break its SBs localized along the z axis and then aligns the other SBs along the z axis similarly to mechanism A. The difference occurs when the SBs have then more difficulty breaking, characterized by a direct denaturation of the beta-strands of CD58.

The ratio of each mechanism in the set of trajectories for CHARMM22 and AMOEBABIO18 has been reported on Table 3. With 18 mechanisms A and only 2 mechanisms B on the 20 trajectories, the CHARMM22 case seems to prefer the first mechanism (18A−2B), whereas the AMOEBABIO18 case seems to prefer the second mechanism (5A−15B). This difference of ratio can thus be related to the energetic gap between the two SDW values. The SDW for the CHARMM22 case is more representative of only one mechanism, although the SDW for the AMOEBABIO18 case is more a mix between the

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

Table 3. SMD Results for the Breaking Process of the CD2−CD58 Complex Using CHARMM22 and AMOEBABIO18 (SMD velocity of 0.05 Å/ps)

Trajectory	CHARMM22	Breaking time (ps)	AMOEBABIO18	Breaking time (ps)
1	√	1030	×	–
2	√	950	√	1070
3	√	920	×	–
4	√	970	×	–
5	×	–	×	–
6	√	1140	×	–
7	√	970	√	1020
8	√	960	×	–
9	√	950	×	–
10	√	1020	×	–
11	×	–	×	–
12	√	980	×	–
13	√	940	×	–
14	√	1070	×	–
15	√	1020	×	–
16	√	970	√	1250
17	√	960	×	–
18	√	940	√	1200
19	√	1010	√	1010
20	√	1070	×	–
	Breaking's (Unbreaking's) probability	Breaking time average (ps)	Breaking's (Unbreaking's) probability	Breaking time average (ps)
	0.88 (0.12)	993	0.25 (0.75)	1110

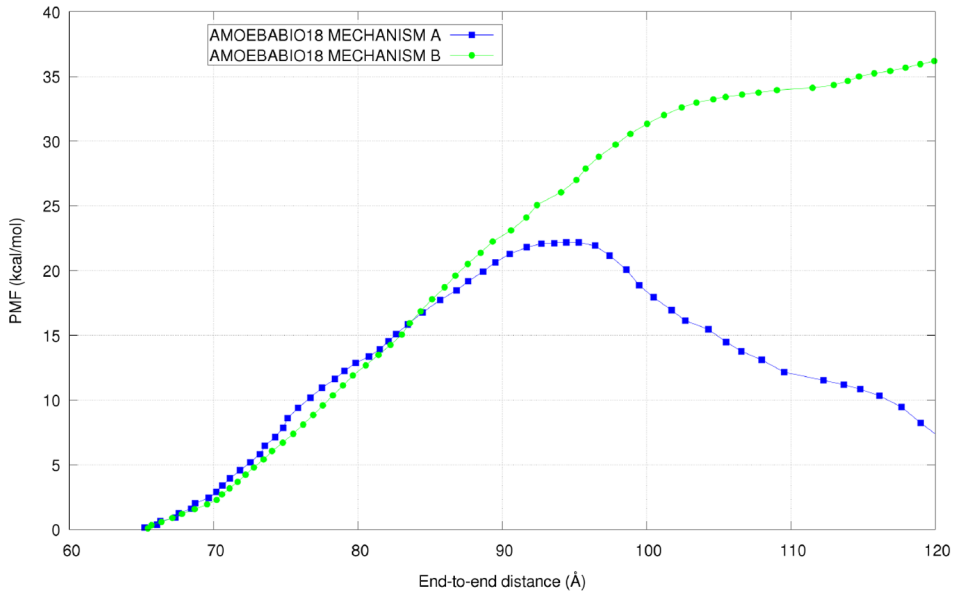


Figure 13. CD2−CD58 potential of mean force calculated with the 15 trajectories of mechanism B of AMOEBABIO18 (green curve) and with the 5 trajectories of mechanism A (blue curve).

two mechanisms A and B. It thus gives full account of a harder energetic convergence. However, if trajectories for the mechanism A (5) and B (15) are treated separately, as represented on Figure 13, it can be seen that they share a similar behavior at the beginning and then diverge. The mechanism A shows a free energy barrier of 22 kcal/mol and the mechanism B of 33 kcal/mol, showing that mechanism A should be preferred compared to mechanism B with a lower free energy barrier than CHARMM22. With an experimental enthalpy of dissociation of the complex evaluated around 19.4 kcal/mol,[90] it encourages us to be more confident about the AMOEBABIO18 results, even if contamination by mechanism B is strongly observed.

This difference in terms of mechanisms was already observed in several works such as the ones of Bayas et al.,[24] Maruthamathu et al.,[91] and Mikulska et al.[92] and directly related to the range of the initial SMD velocities chosen. But in the current case, the SMD velocity is kept constant while the nature of the force field used is changed. Two explanations can be advanced:

- *The water breathing during the pulling process (polarizable vs non-polarizable solvent).* As depicted in the inset of Figure 10, due to the absence of explicit polarization, CHARMM22 water molecules solvate the ionic charges

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

more easily due to a faster reorganization compared to AMOEBABIO18 where relaxation of the polarizable solvent occurs. To support this idea, 20 independent trajectories have been performed in CHARMM22 by increasing the SMD velocity by a factor 10 (0.5 Å/ps). The same difficulty to properly relax the water molecules around the broken SBs as for AMOEBABIO18 was found. In other words, with standard velocities, classical force fields encompass solvent relaxation that are too fast. However, when this relaxation can correctly occur during the five trajectories of the mechanism A in AMOEBABIO18, it has a dramatic impact on the final free energy barrier. This motivates decreasing the SMD velocity when polarizable force fields are used. The simulation time per trajectory will increase, but the number of trajectories necessary to reach convergence will decrease.

- *The strength of beta-strands after the realignment process.* Indeed, when a major part of the SBs is broken, beta-strands of CD58 are directly exposed to the pulling force axis. Beta-strands offer a stronger resistance to the external pulling in the CHARMM22 case than the AMOEBABIO18 one. The idea of a stronger stabilization of beta-strands for the CHARMM non-polarizable force fields could be justified by a recent work of Hazel et al. (2018), where they demonstrate an overstabilization of the beta-strands in case of the CHARMM36 non-polarizable field.[93] Therefore, it could be easier for AMOEBABIO18 to break beta-strands at SMD velocity values where CHARMM22 do not break them easily. Still, to corroborate this idea, we decided to reproduce our SMD study in NAMD and using the CHARMM36 force field. Depicted on Figure 10, results go in the same direction of the CHARMM22 force field with a high preference for mechanism A (19A−1B) and a perfect overlap with CHARMM22 in terms of reorganization of water molecules around the ionic charges of the SBs atoms, showing that SBs prefer to be broken compared to the beta-strands of CD58 for this case. Similar results were also found by using the CHARMM22/CMAP and AMBER99SB non-polarizable force fields. Another study was to take all 20 independent trajectories performed with CHARMM22 with a higher SMD velocity (0.5 Å/ps). They all display mechanism B, with a too high free energy barrier and an SDW in the same range as AMOEBABIO18 (see Table 4). It finally demonstrates that the relaxation of the system is also important during a pulling process and is directly related to the initial SMD velocity chosen, underlying the same conclusions as the previous point for the water breathing.

Table 4. Free Energy Barriers and SDWs Calculated with a Set of 20 Trajectories for Each Force Field (CHARMM22 and AMOEBABIO18) for the CD2−CD58 Complex[a]

force field	SMD velocity (Å/ps)	free energy barrier (kcal/mol)	SDW (kcal/mol)	breaking mechanism
CHARMM22	0.5	89.96	5.76	B
CHARMM22	0.05	30.08	2.38	A
CHARMM36	0.05	35.27	3.81	A
AMOEBABIO18	0.05	28.61	5.40	B

[a]The specific SMD velocity and the breaking mechanism of the salt bridges have been specified for each PMF. The free energy barriers correspond to the free energy obtained when the plateau is reached (so when the CD2−CD58 complex is completely broken).

CONCLUSION AND PERSPECTIVES

We presented our implementation strategy of the Steered Molecular Dynamics (SMD) method within the massively parallel Tinker-HP software. It offers the possibility of sampling rare events for large scale systems using point dipole polarizable force fields such as AMOEBA when other methods such as the Umbrella Sampling are computationally out of reach. We assessed the quality of our new implementation by using both reversible and irreversible SMD velocity regimes to reproduce a PMF along a reference reaction coordinate. Whereas the PMF can directly be obtained by working to pull a measurement on only one SMD trajectory using a slow SMD velocity in the reversible regime, we confirmed that it can also be reproduced by using a fast SMD velocity in the irreversible regime if several trajectories are regrouped through the Jarzynski equality. Averaging over 20 trajectories has been shown to be a reasonable strategy to compute PMFs if one uses the stiff spring approximation and the cumulant expansion method, offering an important gain in simulation time compared to the reversible regime. This methodology was then applied to several benchmark systems using polarizable (AMOEBAPRO13 and AMOEBABIO18) and non-polarizable (CHARMM22/CMAP, AMBER99, AMBER99SB, and OPLS−AA/L) force fields (FFs). Our results show that these two classes of methods are in disagreement when considered in vacuum. Indeed, it can be easily understood that classical FFs are designed and parametrized for the condensed phase only whereas polarizable ones are more transferable thanks to the explicit presence of induced dipoles. As expected, in the condensed phase, FFs present a better convergence when used within a water solvated environment (case of the ubiquitin). Of course, differences also exist in the condensed phase as illustrated by the larger scale CD2−CD58 study. If it is shown that PMFs for non-polarizable (CHARMM22) and polarizable (AMOEBABIO18) force fields provide a close free energy barrier, we have pointed out that they do not predict the same physical mechanisms. Indeed, two different and competitive mechanisms have been characterized. Relaxation of the water molecules around the broken SBs and secondary structures (beta-strands) of the CD58 subpart have been found to play a crucial role in the final unfolding mechanism of the CD2−CD58 complex. While these two features are directly related to the initial chosen SMD velocity, the polarizable AMOEBABIO18 provides a lower and better prediction of the free energy barrier of the main mechanism compared to CHARMM22. On the opposite, the AMOEBA results are strongly contaminated by the second mechanism due to a too high initial SMD velocity. This leads us to the conclusion that one should decrease as much as possible the SMD velocity when using polarizable force fields on large systems to enable the relaxation of the polarizable solvent. Such inclusion of polarization clearly provides an improvement of the quantitativity of the evaluation of the PMFs but, of course, generates an increased computational cost. Nevertheless if the additional cost of the computation of polarization is real, the use of lower velocities for polarizable FFs is not fully an additional burden as such lower velocities also reduce the need of multiple trajectories required by the use of an irreversible regime, therefore diminishing the total cost. Overall, two key differences exist between FF classes. First, it is of importance to allow the

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

solvent structure to relax. We observed that it is clearly dependent on the nature of the force field: classical water relaxes too quickly compared to polarizable water. Moreover, PMF free energy barriers computed with the polarizable AMOEBA always decrease compared to non-polarizable force fields.

Future efforts will continue toward automatic procedures designed to estimate the optimal number of trajectories to ensure a good convergence of the PMFs as a function of the SMD velocity. The presented methodology will also strongly benefit from the engoing Tinker-HP software improvements. Indeed, the initial SMD results presented here were computed using the 1.1 version, and various "acceleration features" were not used. However, our implementation being validated, one can choose to select the fastest setup which will be clearly enhanced in Tinker-HP version 1.2 (and higher), including fast and accurate multi-time step algorithms[94] coupled to noniterative polarization solvers,[95,96] full software vectorization boost,[62] etc. This will lead to a short-term acceleration of, at least, a factor of 2−3 and allow performing larger and more converged SMD simulations with slower SMD velocities as well as larger scale reference Umbrella Sampling simulations to better compare the capabilities of classical and polarizable force fields.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.9b00199.

SI-1: movie of the deca-alanine SMD simulation (AVI)

SI-2: movie of the ubiquitin SMD simulation (AVI)

SI-3: movie of CD2−CD58 SMD simulation (mechanism A) (AVI)

SI-4: movie of CD2−CD58 SMD simulation (mechanism B) (AVI)

■ AUTHOR INFORMATION

Corresponding Author

*(J-P.P.) E-mail: jean-philip.piquemal@sorbonne-universite.fr.

ORCID

Etienne Derat: 0000-0002-8637-2707

Jean-Philip Piquemal: 0000-0001-6615-9426

Funding

This work was made possible thanks to the French state funds managed by the CalSimLab LABEX and the ANR within the Investissements d'Avenir program (reference ANR-11-IDEX-0004-02) and through support of the Direction Génerale de l'Armement (DGA) Maîtrise NRBC of the French Ministry of Defense. Computations have been performed at GENCI on the Occigen machine (CINES, Montpellier, France) except for the CD2−CD58 complex. Such computations were made possible thanks to the GENCI Grand Challenge on the Irène Joliot-Curie supercomputer (TGCC, Bruyères le Châtel, France).

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Antunes, D. A.; Devaurs, D.; Kavraki, L. E. Understanding the challenges of protein flexibility in drug design. *Expert Opin. Drug Discovery* **2015**, *10*, 1301−1313.

(2) Forli, S.; Huey, R.; Pique, M. E.; Sanner, M.; Goodsell, D. S.; Olson, A. J. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* **2016**, *11*, 905−919.

(3) Hernández-Santoyo, A., Tenorio-Barajas, A. Y., Altuzar, V., Vivanco-Cid, H., Mendoza-Barrera, C., Eds. *Protein-Protein and Protein-Ligand Docking, Protein Engineering*; Protein Engineering Tomohisa Ogawa, IntechOpen: 2013. Available from: https://www.intechopen.com/books/protein-engineering-technology-and-application/protein-protein-and-protein-ligand-docking.

(4) Miloshevsky, G. V.; Jordan, P. C. Gating gramicidin channels in lipid bilayers: reaction coordinates and the mechanism of dissociation. *Biophys. J.* **2004**, *86*, 92−104.

(5) Ngo, V. A.; Kim, I.; Allen, T. W.; Noskov, S. Y. Estimation of potentials of mean force from nonequilibrium pulling simulations using both Minh-Adib estimator and Weighted Histogram Analysis Method. *J. Chem. Theory Comput.* **2016**, *12*, 1000−1010.

(6) Wong, K.-Y.; York, D. M. Exact Relation between potential of mean force and free energy profile. *J. Chem. Theory Comput.* **2012**, *8*, 3998−4003.

(7) Nitschke, N.; Atkovska, K.; Hub, J. S. Accelerating potential of mean force calculations for lipid membrane permeation: systeme size, reaction coordinate, solute-solute distance and cutoffs. *J. Chem. Phys.* **2016**, *145*, 125101−1.

(8) Jiang, L.; Gao, Y.; Mao, F.; Liu, Z.; Lai, L. Potential of mean force for protein-protein interaction studies. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 190−196.

(9) Yu, T.; Lee, O.-S.; Schatz, G. C. Steered Molecular Dynamics Studies of the Potential of Mean Force for peptide Amphiphile Self-Assembly into ctlindrical nanofibers. *J. Phys. Chem. A* **2013**, *117*, 7453−7460.

(10) Tavares, F. W.; Bratko, D.; Blanch, H. W.; Prausnitz, J. M. Ion-Specific Effects in the Colloid- Colloid or Protein- Protein Potential of Mean Force: Role of Salt- Macroion van der Waals Interactions. *J. Phys. Chem. B* **2004**, *108*, 9228−9235.

(11) Liu, S.; Zhang, C.; Zhou, H.; Zhou, Y. A physical reference state unifies the structure-derived potential of mean force for protein folding and binding. *Proteins: Struct., Funct., Genet.* **2004**, *56*, 93−101.

(12) Torrie, G. M.; Valleau, J. P. Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chem. Phys. Lett.* **1974**, *28*, 578−581.

(13) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919.

(14) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(15) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.

(16) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

(17) Lagardère, L.; Jolly, L. H.; Lipparini, F.; Aviat, F.; Stamm, B.; Jing, Z. F.; Harger, M.; Torabifard, H.; Cisneros, G. A.; Schnieders, M. J.; Gresh, N.; Maday, Y.; Ren, P. Y.; Ponder, J. W.; Piquemal, J.-P. Tinker-HP: a massively parallel molecular dynamics package for multiscale simulations of large complex systems with advanced point dipole polarizable force fields. *Chem. Sci.* **2018**, *9*, 956−972.

(18) Kästner, J. Umbrella Sampling. *Wiley. Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 932−942.

(19) Kumar, S. K.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for free-energy calculations on Biomolecules I. The method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(20) Izrailev, S., Stepaniants, S., Isralewitz, B., Kosztin, D., Lu, H., Molnar, F., Wriggers, W., Schulten, K., Eds. Steered molecular dynamics. *Computational molecular dynamics: challenges, methods, ideas*; Springer: Berlin, 1997; Chapter 1.

(21) Isralewitz, B.; Gao, M.; Schulten, K. Steered molecular dynamics and mechanical functions of proteins. *Curr. Opin. Struct. Biol.* **2001**, *11*, 224−230.

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

(22) Izrailev, S.; Stepaniants, S.; Balsera, M.; Oono, Y.; Schulten, K. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* **1997**, *72*, 1568−1581.
(23) Gao, M.; Wilmanns, M.; Schulten, K. Steered molecular dynamics studies of titin I1 domain unfolding. *Biophys. J.* **2002**, *83*, 3435−3445.
(24) Bayas, M. V.; Schulten, K.; Leckband, D. Forced detachment of the CD2-CD58 complex. *Biophys. J.* **2003**, *84*, 2223−2233.
(25) Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690.
(26) Jarzynski, C. Equilibrium free-energy differences from non-equilibrium measurements: A master-equation approach. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1997**, *56*, 5018.
(27) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. Free energy calculation from steered molecular dynamics simulations using Jarzynski equality. *J. Chem. Phys.* **2003**, *119*, 3559−3566.
(28) Park, S.; Schulten, K. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* **2004**, *120*, 5946−5961.
(29) Jensen, M. Ø.; Park, S.; Tajkhorshid, E.; Schulten, K. Energetics of glycerol conduction through aquaglyceroporin GlpF. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 6731−6736.
(30) Hummer, G. fast-growth thermodynamic integration: Error and efficiency analysis. *J. Chem. Phys.* **2001**, *114*, 7330−7337.
(31) Ozer, G.; Valeev, E. F.; Quirk, S.; Hernandez, R. Adaptative Steered Molecular Dynamics of the long-distance unfolding of neuropeptide Y. *J. Chem. Theory Comput.* **2010**, *6*, 3026−3038.
(32) Gu, J.; Li, H.; Wang, X. A Self-Adaptative Steered Molecular Dynamic method based on minimization of stretching force reveals the binding affinity of protein-ligand complexes. *Molecules* **2015**, *20*, 19236−19251.
(33) Crooks, G. E. Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems. *J. Stat. Phys.* **1998**, *90*, 1481−1487.
(34) Crooks, G. E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1999**, *60*, 2721−2726.
(35) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.
(36) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory. Comput.* **2008**, *4*, 435−447.
(37) Lamoureux, G.; Mackerell, A. D.; Roux, B. A simple polarizable model of water based on classical drude oscillators. *J. Chem. Phys.* **2003**, *119*, 5185−5197.
(38) Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; Mackerell, A. D. J. Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2013**, *9*, 5430−5449.
(39) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. Anisotropic, polarizable molecular mechanics studies of inter-, intramolecular interactions, and ligand-macromolecule complexes. A bottom-up strategy. *J. Chem. Theory Comput.* **2007**, *3*, 1960−1986.
(40) Shi, Y.; Ren, P.; Schnieders, M.; Piquemal, J.-P. In *Polarizable Force Fields for Biomolecular Modeling*; Parrill, A. L., Lipkowitz, K. B., Eds.; Wiley: 2015; Vol. *28*, pp 51−84.
(41) Piquemal, J.-P.; Jordan, K. D. Preface: Special Topic: From Quantum Mechanics to Force Fields. *J. Chem. Phys.* **2017**, *147*, 161401.
(42) Ren, P.; Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003**, *107*, 5933−5947.
(43) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. The Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. *J. Chem. Theory. Comput.* **2013**, *9*, 4046−4063.
(44) Zhang, C.; Lu, C.; Jing, Z.; Wu, C.; Piquemal, J.-P.; Ponder, J. W.; Ren, P. AMOEBA Polarizable Atomic Multipole Force Field for Nucleic Acids. *J. Chem. Theory Comput.* **2018**, *14*, 2084−2108.
(45) Sandoval-Perez, A.; Pluhackova, K.; Böckmann, R. A. Critical Comparison of Biomembrane Force Fields: Protein-Lipid Interactions at the Membrane Interface. *J. Chem. Theory Comput.* **2017**, *13*, 2310−2321.
(46) Guvench, O., MacKerell, A. D. J., Eds. *Molecular modeling of proteins: Comparison of protein force fields for molecular dynamics simulations*; Humana Press: 2008.
(47) Price, D. J.; Brooks, C. L. Modern Protein Force Fields Behave Comparably in Molecular Dynamics Simulations. *J. Comput. Chem.* **2002**, *23*, 1045−1057.
(48) Petrov, D.; Zagrovic, B. Are Current Atomistic Force Fields Accurate Enough to Study Proteins in Crowded Environments? *PLoS Comput. Biol.* **2014**, *10*, e1003638−12.
(49) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. Extremely precises free energy calculations of amino acids side chain analogs: comparison of common molecular mechanics force fields for proteins. *J. Chem. Phys.* **2003**, *119*, 5740−5761.
(50) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Systematic validation of protein force fields against experimental data. *PLoS One* **2012**, *7*, e32131.
(51) Cino, E. A.; Choy, W.-Y.; Karttunen, M. Comparison of secondary structure formation using 10 different force fields micro-second molecular dynamics simulations. *J. Chem. Theory Comput.* **2012**, *8*, 2725−2740.
(52) Beauchamp, K. A.; Lin, Y.-S.; Das, R.; Pande, V. S. Are protein force fields getting better ? A systematic benchmark on 524 diverse NMR measurements. *J. Chem. Theory Comput.* **2012**, *8*, 1409−1414.
(53) Lange, O. F.; van der Spoel, D.; de Groot, B. L. Scrutunizing molecular mechanics force fields on the submicrosecond timescale with NMR data. *Biophys. J.* **2010**, *99*, 647−655.
(54) Best, R. B.; Buchete, N.-V.; Hummer, G. Are current molecular dynamics force fields too helical? *Biophys. J.* **2008**, *95*, L07−L09.
(55) Kollman, P. Free energy calculations: Applications to chemical and biochemical phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.
(56) Hansen, N.; van Gunsteren, W. Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.* **2014**, *10*, 2632−2647.
(57) Kirkwood, J. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **1935**, *3*, 300−313.
(58) Rackers, J. A.; Wang, Z.; Lu, C.; Laury, M. L.; Lagardère, L.; Schnieders, M. J.; Piquemal, J.-P.; Ren, P.; Ponder, J. W. Tinker 8: Software Tools for Molecular Design. *J. Chem. Theory Comput.* **2018**, *14*, 5273−5289.
(59) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(60) Lipparini, F.; Lagardère, L.; Stamm, B.; Cancès, E.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J.-P. Scalable Evaluation of Polarization Energy and Associated Forces in Polarizable Molecular Dynamics: I. Toward Massively Parallel Direct Space Computations. *J. Chem. Theory Comput.* **2014**, *10*, 1638−1651.
(61) Lagardère, L.; Lipparini, F.; Polack, È.; Stamm, B.; Cancès, É.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J.-P. Scalable Evaluation of Polarization Energy and Associated Forces in Polarizable Molecular Dynamics: II. Toward Massively Parallel Computations Using Smooth Particle Mesh Ewald. *J. Chem. Theory Comput.* **2015**, *11*, 2589−2599.
(62) Jolly, L. H.; Duran, A.; Ponder, J. W.; Ren, P. Y.; Lagardère, L.; Piquemal, J.-P. Raising the Performance of the Tinker-HP Molecular Modeling Package on Intelas HPC Architectures: a Living Review [Article v1.0]. **2019**, submitted.
(63) Swope, W.; Andersen, H.; Berens, P.; Wilson, K. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982**, *76*, 637.
(64) Tuckerman, M.; Berne, B. J.; Martyna, G. J. Reversible multiple time scale molecular dynamics. *J. Phys. Chem. B* **1992**, *97*, 1990.

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

(65) Leimkuhler, B.; Matthews, C. Efficient molecular dynamics using geodesic integration and solvent-solute splitting. *Proc. R. Soc. London, Ser. A* **2016**, *472*, 20160138.
(66) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity-rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
(67) Grossfield, A. WHAM: the weighted histogram analysis method. http://membrane.urmc.rochester.edu/content/wham.
(68) Hazel, A. J.; Chipot, C.; Gumbart, J. C. Thermodynamics of Deca-alanine Folding in Water. *J. Chem. Theory Comput.* **2014**, *10*, 2836−2844.
(69) Aqvist, J. Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. *J. Phys. Chem.* **1990**, *94*, 8021−8024.
(70) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(71) Ross, W. S.; Hardin, C. C. Ion-Induced Stabilization of the G-DNA Quadruplex: Free Energy Perturbation Studies. *J. Am. Chem. Soc.* **1994**, *116*, 6070−6080.
(72) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calcluating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(73) Moyna, G.; Williams, H. J.; Nachman, R. J.; Scott, A. I. Conformation in Solution and Dynamics of a Structurally Constrained Linear Insect Kinin Pentapeptide Analogue. *Biopolymers* **1999**, *49*, 403−413.
(74) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712−725.
(75) Foloppe, N.; MacKerell, A. D., Jr. All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem.* **2000**, *21*, 86−104.
(76) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(77) MacKerell, A. D. J.; Feig, M.; Brooks, C. L. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(78) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(79) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(80) Liu, C.; Ponder, J. W.; Marshall, G. R. Helix stability of oligoglycine, oligoalanine, and oligo-β-alanine dodecamers reflected by hydrogen-bond persistence. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 3043−3061.
(81) Sindelar, C. V.; Hendsch, Z. S.; Tidor, B. Effects of salt bridges on protein structure and design. *Protein Sci.* **1998**, *7*, 1898−1914.
(82) Bosshard, H. R.; Marti, D. N.; Jelesarov, I. Protein stabilization by salt bridges: concepts,experimental approaches and clarification ofsome misunderstandings. *J. Mol. Recognit.* **2004**, *17*, 1−16.
(83) Leontyev, I.; Stuchebrukhov, A. Accounting for electronic polarization in non-polarizable force fields. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2613−2614.
(84) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization? *Biophys. J.* **2011**, *100*, L47−L49.
(85) Debiec, K. T.; Gronenborn, A. M.; Chong, L. T. Evaluating the Strength of Salt Bridges: A Comparison of Current Biomolecular Force Fields. *J. Phys. Chem. B* **2014**, *118*, 6561−6569.
(86) Debiec, K. T.; Cerutti, D. S.; Baker, L. R.; Gronenborn, A. M.; Case, D. A.; Chong, L. T. Further along the Road Less Traveled: AMBER ff15ipq, an Original Protein Force Field Built on a Self-Consistent Physical Model. *J. Chem. Theory Comput.* **2016**, *12*, 3926−3947.
(87) Houriez, C.; Vallet, V.; Réal, F.; Meot-Ner, M.; Masella, M. Organic ion association in aqueous phase and ab initio-based force fields: The case of carboxylate/ammonium salts. *J. Chem. Phys.* **2017**, *147*, 161720.
(88) Ahmed, M. C.; Papaleo, E.; Lindorff-Larsen, K. How well do force fields capture the strength of salt bridges in proteins? *PeerJ* **2018**, *6*, e4967−17.
(89) Ngo, V. A.; Fanning, J. K.; Noskov, S. Y. Comparative Analysis of Protein Hydration from MD simulations with Additive and Polarizable Force Fields. *Adv. Theory. Sim.* **2019**, *2*, 1800106.
(90) Bayas, M. V.; Kearney, A.; Avramovic, A.; van der Merwe, P. A.; Leckband, D. E. Impact of Salt Bridges on the Binding and Adhesion of Human CD2 and CD58. *J. Bio. Chem.* **2007**, *282*, 5589.
(91) Maruthamuthu, V.; Schulten, K.; Leckband, D. Elasticity and Rupture of a Multi-Domain Neural Cell Adhesion Molecule Complex. *Biophys. J.* **2009**, *96*, 3005−3014.
(92) Mikulska, K.; Strzelecki, J.; Nowak, W. Nanomechanics of β-rich proteins related to neuronal disorders studied by AFM, all-atom and coarse-grained MD methods. *J. Mol. Model.* **2014**, *20*, 2144.
(93) Hazel, A. J.; Walters, E. T.; Rowley, C. N.; Gumbart, J. C. Folding free energy landscapes of β-sheets with non-polarizable and polarizable CHARMM force fields. *J. Chem. Phys.* **2018**, *149*, 072317.
(94) Lagardère, L.; Aviat, F.; Piquemal, J.-P. Pushing the limits of Multiple-Timestep Strategies for Polarizable Point Dipole Molecular Dynamics. *J. Phys. Chem. Lett.* **2019**, *10*, 2593−2599.
(95) Aviat, F.; Levitt, A.; Stamm, B.; Maday, Y.; Ren, P.; Ponder, J. W.; Lagardère, L.; Piquemal, J. Truncated Conjugate Gradient (TCG): an optimal strategy for the analytical evaluation of the many-body polarization energy and forces in molecular simulations. *J. Chem. Theory. Comput.* **2017**, *13*, 180−190.
(96) Aviat, F.; Lagardère, L.; Piquemal, J.-P. The Truncated Conjugate Gradient (TCG), a Non-iterative/Fixed-cost Strategy for Computing Polarization in Molecular Dynamics: Fast Evaluation of Analytical Forces. *J. Chem. Phys.* **2017**, *147*, 161724.

DOI: 10.1021/acs.jctc.9b00199
J. Chem. Theory Comput. 2019, 15, 3694−3709

JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Erratum

Correction to "Massively Parallel Implementation of Steered Molecular Dynamics in Tinker-HP: Comparisons of Polarizable and Nonpolarizable Simulations of Realistic Systems"

Frédéric Célerse, Louis Lagardère, Etienne Derat, and Jean-Philip Piquemal*

J. Chem. Theory Comput. **2019**, *15*, 6, 3694−3709. DOI: 10.1021/acs.jctc.9b00199

Cite This: *J. Chem. Theory Comput.* 2021, 17, 3235−3236 Read Online

ACCESS | Metrics & More | Article Recommendations

While performing routine tests using the newly released GPUs version[1] of the Tinker-HP package,[2] we noticed that there was an error in the potential of mean force (PMF) calculation procedure described in our paper. Indeed, a scripting error led to the introduction of wrong SMD vector components for Figure 9b. This does not change the discussion concerning the differences between forces fields and the conclusion of this section of the paper. Nevertheless, in Figure 1 we provide here the corrected PMF for the stretching process of ubiquitin's β strand computed using various force fields since the new results open a discussion about the convergence of the SMD simulations.

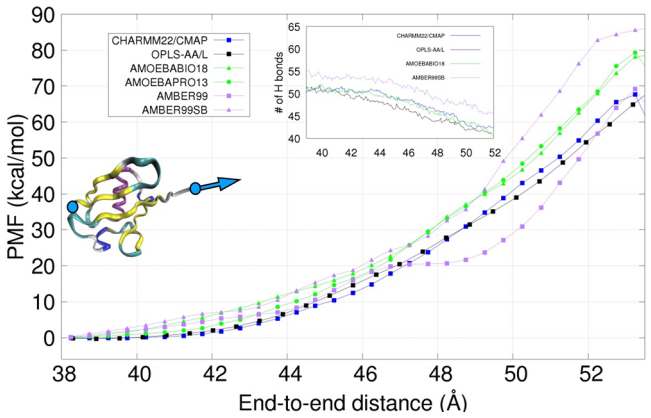


Figure 1. Corrected Figure 9b.

As Table 2 of the paper is directly associated with Figure 9, the corrected relative values for Table 2 are also provided here in Table 1.

Analyzing these new numbers, we can stress that the relatively high values of the standard deviation work (SDW) could tend to show that our SMD calculations are not fully converged. Although SDWs are only partial indicators of convergence, we decided to perform further simulations. Indeed, a previous work performed by Pai-Chi et al.[3] using umbrella sampling (US) coupled to the CHARMM22/CMAP force field showed that they obtained a smaller free energy barrier of approximatively 31 kcal/mol. In this direction, we decided to decrease the SMD velocity by a factor 10, from 0.01 to 0.001 Å/ps. Thanks to GPUs,[1] additional PMFs have been calculated for the nonpolarizable force field CHARMM22/CMAP and the polarizable force field AMOEBABIO18. We also completed these results with additional US simulations to check the convergence of our new SMD results. The setup of the US simulations (time step, integrator, thermostat, ...) was similar to that of the SMD simulations. The new results are depicted in Figure 2.

Table 1. Corrected Table 2

force field	free energy barrier (kcal/mol)	SDW (kcal/mol)
AMBER99	79.2	5.93
AMBER99SB	82.1	5.46
OPLS-AA/L	71.0	6.35
CHARMM22/CMAP	68.9	4.68
AMOEBAPRO13	79.9	4.89
AMOEBABIO18	78.8	5.00

With this update, we observe that we obtain the same free energy barrier as the one previously obtained by Pai-Chi et al.[3] It is also in good agreement with our respective US results. We also observe that the free energy barrier obtained with AMOEBABIO18 is in agreement with its corresponding PMF obtained using US simulations. Furthermore, it is worth noting that the free energy difference between CHARMM22/CMAP and AMOEBABIO18 amounts to 10 kcal/mol, which is similar to the PMF obtained with a faster SMD velocity (i.e., 0.01 Å/ps).

We can then conclude, on the basis of the literature and on our US simulations that we reached a satisfactory convergence for the ubiquitin system with both CHARMM22/CMAP and AMOEBABIO18 using a slower SMD velocity (0.001 Å/ps). Since the free energy differences between force fields are not impacted by the use of two different SMD velocities (and have

Published: April 28, 2021

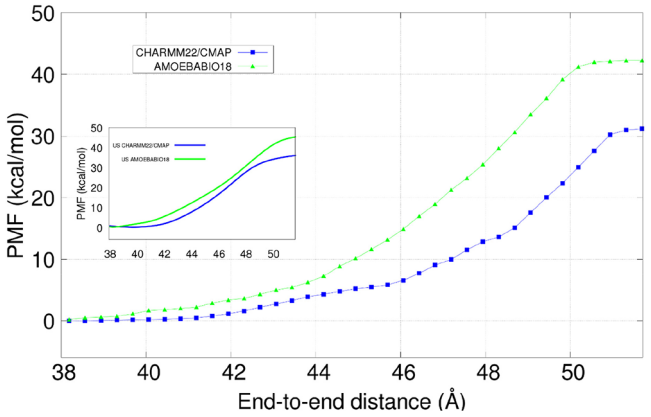


Figure 2. Potential of mean force computations using SMD and umbrella sampling for the ubiquitin protein system (PDB: 1UBQ) using the polarizable AMOEBA and CHARMM22/CMAP force fields. US results are enclosed in the left inset.

been compared to umbrella sampling computations), the conclusions concerning the comparisons between nonpolarizable and polarizable force fields discussed our paper are not affected by the bias induced by the initial excessive SMD velocity.

REFERENCES

(1) Adjoua, O.; Lagardère, L.; Jolly, L.-H.; Durocher, Arnaud; Wang, Z.; Very, T.; Dupays, I.; Jaffrelot Inizan, T.; Célerse, F.; Ren, P.; Ponder, J.; Piquemal, J.-P. Tinker-HP: Accelerating Molecular Dynamics Simulations of Large Complex Systems with Advanced Point Dipole Polarizable Force Fields using GPUs and Multi-GPUs systems. *J. Chem. Theory Comput.* **2021**, *17* (4), 2034−2053.

(2) Lagardère, L.; Jolly, L.-H.; Lipparini, F.; Aviat, F.; Stamm, B.; Jing, Z. F.; Harger, M.; Torabifard, H.; Cisneros, G. A.; Schnieders, M. J.; Gresh, N.; Maday, Y.; Ren, P.; Ponder, J. W.; Piquemal, J.-P. Tinker-HP: a Massively Parallel Molecular Dynamics Package for Multiscale Simulations of Large Complex Systems with Advanced Polarizable Force Fields. *Chem. Sci.* **2018**, *9*, 956−972.

(3) Li, P.-C.; Makarov, D. E. Simulation of the mechanical unfolding of ubiquitin: probing different unfolding reaction coordinates by changing the pulling geometry. *J. Chem. Phys.* **2004**, *121* (10), 4826−4832.

https://doi.org/10.1021/acs.jctc.1c00405
J. Chem. Theory Comput. 2021, 17, 3235−3236